Am. J. Hum. Genet. 59:437–444, 1996

mtDNA Sequence Diversity in Africa

Elizabeth Watson,[1] Karin Bauer,[1] Rashid Aman,[2] Gunter Weiss,[1] Arndt von Haeseler,[1] and Svante Pääbo[1]

[1]Institute of Zoology, University of Munich, Munich; and [2]Department of Molecular Genetics, National Museums of Kenya, Nairobi

Summary

mtDNA sequences were determined from 241 individuals from nine ethnic groups in Africa. When they were compared with published data from other groups, it was found that the !Kung, Mbuti, and Biaka show on the order of 10 times more sequence differences between the three groups, as well as between those and the other groups (the Fulbe, Hausa, Tuareg, Songhai, Kanuri, Yoruba, Mandenka, Somali, Tukana, and Kikuyu), than these other groups do between one other. Furthermore, the pairwise sequence distributions, patterns of coalescence events, and numbers of variable positions relative to the mean sequence difference indicate that the former three groups have been of constant size over time, whereas the latter have expanded in size. We suggest that this reflects subsistence patterns in that the populations that have expanded in size are food producers whereas those that have not are hunters and gatherers.

Introduction

Much interest has focused on mtDNA variation as a tool to unravel the past of modern humans. The main reason for this is that the absence of recombination and the high evolutionary rate of the mitochondrial genome offer a resolving power unparalleled by any other single genetic locus in the human genome. Studies of mitochondrial genetic diversity on a worldwide scale initially used restriction-enzyme polymorphism (Cann et al. 1987). Later this was followed by studies of DNA sequences of the mitochondrial control region (Vigilant et al. 1991). Both data sets have been used to infer an African origin for the human mitochondrial gene pool and thus, by inference, for modern humans. This conclusion has been contested (e.g., see Hedges et al. 1992; Maddison et al. 1992; Templeton 1992, 1993) as well as supported (e.g., see Horai et al. 1995; Penny et al. 1995; Zischler et al. 1995). Over the past few years, large numbers of sequences, particularly of the hypervariable region I (HVR I) of the mitochondrial control region, have been determined from individuals representing many regions of the world (DiRienzo and Wilson 1991; Ward et al. 1991, 1993; Piercy et al. 1993; Shields et al. 1993; Torroni et al. 1993*a*, 1993*b*; Pult et al. 1994; Santos et al. 1994; Batista et al. 1995; Bertranpetit et al. 1995; Kolman et al. 1995; Mountain et al. 1995; Sajantila et al. 1995). However, apart from the six African groups originally studied (Vigilant et al. 1991), HVR I sequences from only one additional African group have been published (Graven et al. 1995). In order to survey the mitochondrial sequence diversity in Africa, we have determined HVR I sequences from a total of 241 individuals from nine ethnic groups in Africa.

Received January 26, 1996; accepted for publication May 17, 1996.

Address for correspondence and reprints: Dr. Svante Pääbo, Institute of Zoology, University of Munich, P.O. Box 202136, D-80021 München, Germany. E-mail: paabo@zi.biologie.uni-muenchen.de

0002-9297/96/5902-0021$02.00

Subjects, Material, and Methods

Blood samples were collected from unrelated individuals at hospitals and rural medical clinics in Kenya, Nigeria, and Niger. The ethnicity and place of birth of the individuals, as well as of their parents and, whenever possible, of their maternal and paternal grandparents, were noted. All samples were obtained after informed consent under forms approved by the Massey University Ethics Committee was obtained.

Two aliquots of 0.75 ml of whole blood were drawn from each individual and were suspended in 0.75 ml of 100 mM Tris, 100 mM EDTA, and 1% SDS, and samples were stored at ambient temperatures. To extract DNA, 400 ml of the blood/buffer solution was added to 200 ml of NaCl (100 mM), DTT (80 mg/ml), and proteinase K (10 mg/ml) and was incubated at 37°C for 3 h. After phenol and chloroform extractions, the DNA was concentrated in Centricon tubes (Amicon). PCR amplification of the HVR I of the mitochondrial control region was performed by using the primers H16498 5′-CCT GAA GTA GGA ACC AGA TG-3′ (biotinylated) and L15926 5′-TCA AAG CTT ACA CCA GTC TTG TAA ACC-3′. The strands were separated by using Dynabeads (Dynal) and were sequenced by using Sequenase 2.0 (United States Biochemicals) with fluorescent primers H16401 5′-TGA TTT CAC GGA GGA TGG TG-3′ and L15997 5′-CAC CAT TAG CAC CCA AAG CT-3′ on an automated laser fluorescence sequencer (A.L.F.;

Pharmacia). The sequences were determined from position 16024 to position 16383 (Anderson et al. 1981) and were aligned by using ESEE (Cabot 1993).

The 241 new sequences determined were aligned with published control-region sequences. After exclusion of published sequences with one or more undetermined position, 146 previously published (Vigilant et al. 1991; Graven et al. 1995) African sequences remained. These stemmed from 87 Mandenka, 19 !Kung, 17 Biaka, 11 Mbuti, and 12 Yoruba. The Hadza and Herero samples were excluded, since all sequences contained undetermined positions. Pairwise differences within and between the population samples were calculated by using an unpublished program by A.v.H. Genetic distances between two populations (*DS*) were calculated as follows: $DS = \max\{0, D_{12} - (D_1 + D_2)/2\}$, where D_{12} mean pairwise difference between populations 1 and 2 and D_1 and D_2 mean pairwise difference within populations 1 and 2, respectively (Nei 1987). These distance values were used to construct a neighbor-joining tree (Saitou and Nei 1987), as implemented in PHYLIP 3.5 (Felsenstein 1990).

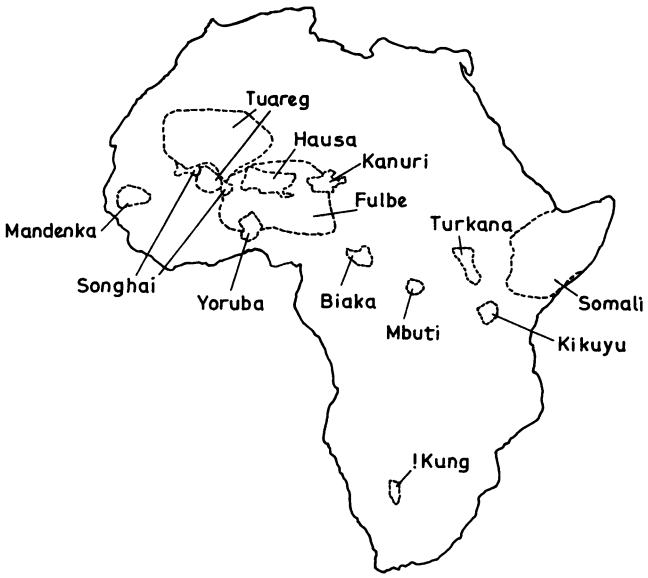

Figure 1 Map of Africa. Boundaries of groups are adapted from Murdock (1959). Dashed lines indicate a wide distribution of an ethnic group, often mixed with other groups. For the Fulbe, which have a larger range, the area indicated represents the region from which the samples analyzed derive.

Results

DNA Sequences

DNA sequences of the HVR I of the mitochondrial control region were determined from 27 Somali, 37 Turkana, and 25 Kikuyu individuals in eastern Africa and from 26 Tuareg, 14 Kanuri, 20 Hausa, 61 Fulbe, 21 Yoruba, and 10 Songhai individuals in western Africa (fig. 1 and table 1). Among the 341 homologous positions determined, 126 positions showed nucleotide differences. At 19 of these positions, transversions were observed. In addition, length variation occurred in and adjacent to a region that in some individuals is composed of 10 consecutive CG base pairs. In total, the sequence differences defined 174 lineages.

In order to obtain an overview of the mitochondrial sequence diversity in Africa, the 241 newly determined sequences and the 148 previously published ones (Vigilant et al. 1991; Graven et al. 1995) were aligned. Among the 389 sequences analyzed, 232 different lineages existed. The mean pairwise difference among all sequences was 8.46 ± 3.36, and the maximum number of observed differences was 21.

African mtDNA Diversity

Table 2 shows the mean pairwise sequence differences within and between the 13 African populations. The within-group sequence diversity of 12 of the groups varies between 6.2 and 10.6, whereas the !Kung have a mean of only 2.9. Figure 2 shows the distributions of pairwise sequence differences within the groups. Ten populations show distributions that are similar to each other in that they have relatively few (0.4%–4.5%) identical sequence pairs and tend to have a bell-shaped distribution with a mode located between 6 and 10 differences. In contrast, three groups—the !Kung, Mbuti, and Biaka (formerly known as Eastern and Western Pygmies, respectively)—have relatively many (11%–31%) identical sequence pairs and show "bumpy," or "rugged," distributions, where no tendency for a defined peak can be discerned. In general, the sample sizes of the three latter groups are smaller (11–19) than those (10–87) of the other groups studied. However, it is unlikely that the bumpy distributions of pairwise sequence in the former groups are due simply to small sample size, because the other groups with small sample sizes (Songhai [10], Kanuri [14], and Hausa [20]) show a clear tendency toward bell-shaped distributions. Furthermore, when the groups with bumpy distributions are pooled, they still show a bumpy distribution, whereas, when other groups with small sample sizes are combined, they show bell-shaped distributions even more clearly than when they are analyzed individually.

When the between-group sequence differences are analyzed (table 2), the mean pairwise differences among the 10 groups that have bell-shaped distributions of sequence differences vary between 6.6 and 10.6, with an average genetic distance corrected for intrapopulation differences (i.e., *DS*) of 0.32. The mean pairwise sequence differences among the three groups with bumpy distributions vary between 8.8 and 9.2, with a mean genetic distance of 3.1. Thus, when distances between

Table 1

Summary of Relevant Information for 13 African Populations

Population	Lifestyle[a]	Population Size (per 1,000)	Sample Size	Mean Pairwise Difference	No. of Variable Sites
!Kung	FC	6	19	2.9	16
Biaka	FC	30	17	8.1	17
Mbuti	FC	30	11	6.4	11
Kikuyu	FP	4,000	25	8.5	49
Somali	FP	2,000	27	7.7	45
Turkana	FP	200	37	10.7	63
Fulbe	FP	8,000	61	7.1	46
Hausa	FP	12,000	20	6.2	33
Songhai	FP	528	10	9.2	31
Tuareg	FP	360	26	7.1	41
Yoruba	FP	15,000	33	7.7	47
Kanuri	FP	1,000	14	7.5	35
Mandenka	FP	800	87	6.6	45

[a] FC = food collecting (i.e., hunter-gatherer); and FP = food production (i.e., pastorialists and agriculturists).

populations are corrected for intragroup diversity, the diversity between the latter groups is 10 times larger than that between the former groups. Furthermore, the pairwise differences between groups with bell-shaped and rugged distributions vary between 9.4 and 10.7, with an average genetic distance of 3.5. Thus, the two types of groups show substantial genetic differentiation between them.

The differentiation between groups with rugged and bell-shaped distributions of pairwise sequence differences is reflected in a neighbor-joining tree (fig. 3). Although the tree is not very stable, the former groups (!Kung, Mbuti, and Biaka) diverge first. Furthermore, the geographic distribution of the populations seems to be reasonably well reflected in the tree in that the populations from western Africa form a cluster. The short branches separating these groups indicate the lack of genetic differentiation among the western African populations. This is different from the eastern African populations and the Biaka, Mbuti, and !Kung, which show a clear separation from each other and from the remaining populations, as indicated by the long internal branches.

Demographic History of the Populations

Populations that have been of constant size over a long time tend to have bumpy distributions, whereas populations that are exponentially growing or have sud-

Table 2

Mean Pairwise Differences within (on the Diagonal, underlined) and between (above the Diagonal) Populations, and Genetic Distances (×100) (below the Diagonal) between Populations

	Ful	Hau	Tua	Son	Kan	Yor	Man	Som	Tur	Kik	Mbu	Bia	!Ku
Fulbe	7.1	6.7	7.2	8.2	7.2	7.4	7.0	7.7	9.8	8.3	10.1	11.5	10.1
Hausa	3	6.2	6.6	7.7	6.7	7.0	6.7	7.2	9.3	7.8	9.5	10.9	9.5
Tuareg	6	0	7.1	8.1	7.1	7.4	7.1	7.5	9.6	8.1	9.7	11.1	9.6
Songhai	4	0	0	9.2	8.4	8.5	8.2	8.9	10.6	9.3	10.5	11.6	10.7
Kanuri	0	0	0	6	7.5	7.5	7.2	7.7	9.7	8.2	9.7	11.7	10.1
Yoruba	1	2	0	2	0	7.7	7.3	7.9	9.7	8.3	9.8	11.0	9.4
Mandenka	16	28	28	32	15	14	6.6	7.3	9.7	8.3	10.1	11.4	10.1
Somali	32	25	9	42	10	25	71	7.7	9.8	8.4	10.1	11.7	9.8
Turkana	93	88	74	68	67	55	112	63	10.7	9.7	10.1	11.9	9.9
Kikuyu	42	43	23	35	17	16	66	21	0	8.7	9.4	11.3	9.4
Mbuti	323	328	300	274	281	278	362	296	163	189	6.4	10.8	8.8
Biaka	391	376	346	293	394	307	401	379	249	294	356	8.1	8.8
!Kung	508	503	464	466	493	413	541	447	316	362	419	336	2.9

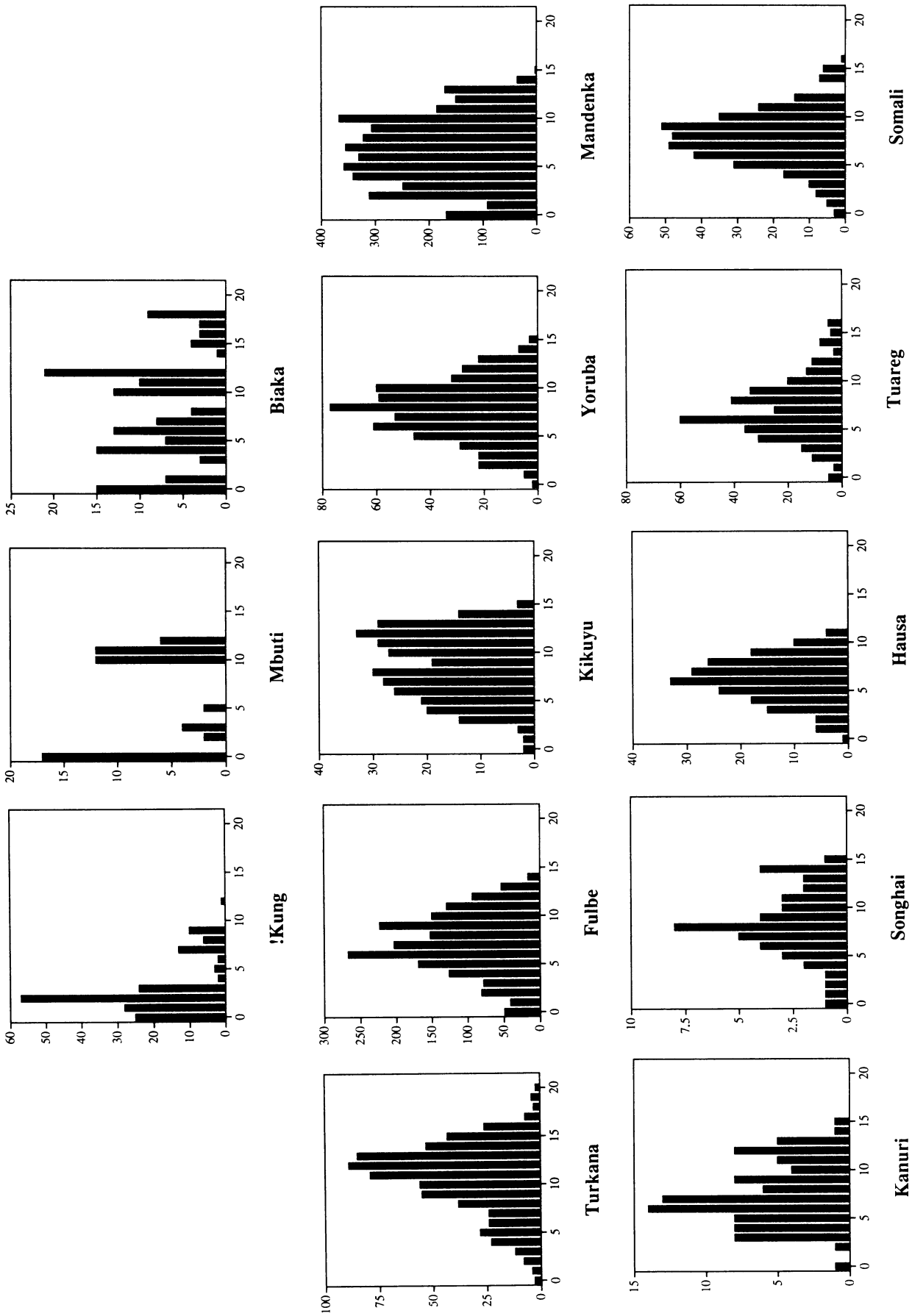

Figure 2 Distributions of pairwise sequence differences within 13 African populations. The abscissa gives the number of differences in a pair, and the ordinate gives the relative number of pairs.

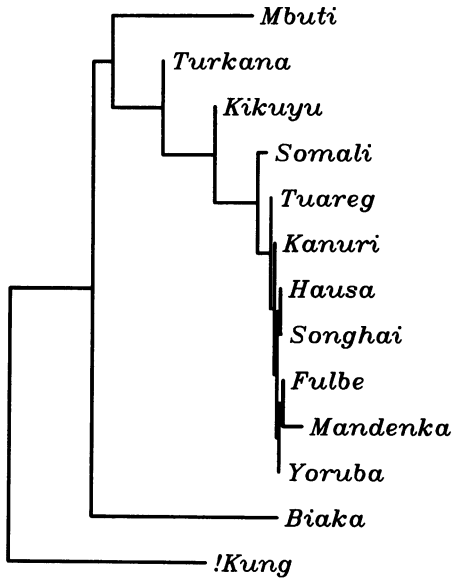

Figure 3 Neighbor-joining tree based on the genetic differences between the African populations (see table 2). The tree was rooted by using the genetic differences between the 13 populations and 60 chimpanzee control-region sequences. Negative branch lengths were set to zero.

denly expanded in the past tend to have distributions of differences that are more smooth and bell shaped (Slatkin and Hudson 1991; Rogers and Harpending 1992). It is therefore tempting to suggest that the difference in distributions between the Biaka, Mbuti, and !Kung and the other populations studied reflects a difference in their demographic history. However, evolutionary scenarios other than the one outlined above may lead to bell-shaped and bumpy distributions (Marjoram and Donnelly 1994).

Fortunately, Nee et al. (1995) have recently developed a graphical method that uses the relative depths of coalescence events in molecular phylogenies to distinguish between populations that have been of constant size and those that have been growing. For this analysis, the sequences from the Biaka, Mbuti, and !Kung were pooled, whereas the sequences from other populations were analyzed separately. In figure 4, the number of lineages in phylogenies of mitochondrial sequences found within the populations are plotted against "time," expressed as transformed amounts of sequence difference (Nee et al. 1995), both under the assumption of constant population size and under the assumption of exponential growth. For !Kung, Biaka, and Mbuti a reasonable fit to a straight line is seen when constant population size is assumed, whereas the fit to a straight line is slightly worse when exponential growth is assumed. In contrast, all other groups display patterns consistent with exponential growth. The Fulbe serve as a representative example in figure 4. Thus, these analyses support the notion that the !Kung, Biaka, and Mbuti have been of constant size over the time covered by the coalescence for the HVR I sequences, whereas the other groups have experienced an increase in population size.

The genealogies generated by the coalescent process depend strongly on the history of a population. Whereas samples from a population of constant size produce a genealogy in which the lengths of branches between coalescent events increase progressively going back in time, the shape of a coalescent tree formed under the assumption of exponential growth is more starlike. This difference in the genealogies results in different ratios of numbers of variable positions to mean pairwise sequence differences. In order to test the hypothesis that the !Kung, Biaka, and Mbuti have been of roughly constant size whereas other African groups have been growing, the mean pairwise sequence differences and the number of variable sites for the populations studied were calculated (table 1). It can be seen that for the !Kung, Biaka, and Mbuti the number of variable sites is always smaller than those for the remaining populations. If a population is of constant size, coalescence theory provides two methods to estimate the parameter θ, defined as $2N\mu$, where N is the effective female population size and μ is

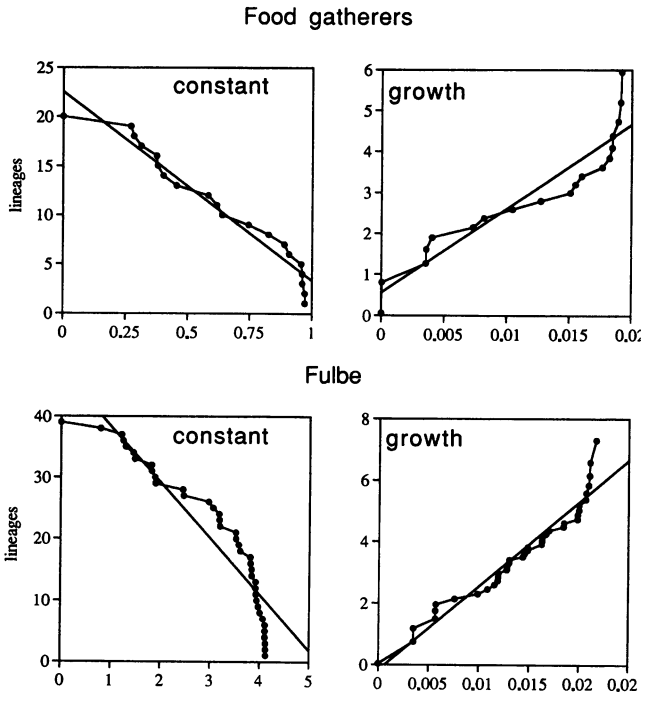

Figure 4 Population plots for food-gathering populations (!Kung, Mbuti, and Biaka) and a food-producing population (Fulbe). The plots on the left display the relationship between times to coalescence and numbers of lineages present under the assumption of constant population sizes. The plots on the right show the decrease of the number of lineages under the "epidemic" transformation, under the assumption of population growth (Nee et al. 1995).

the mutation rate per sequence and generation. First, θ can be estimated from the mean pairwise sequence difference, $\bar{d}$, i.e., $\hat{\theta} = \bar{d}$. Second, θ can be estimated from the number of variable sites S, as $\hat{\theta} = S/\Sigma_{i=1}^{n-1}(1/i)$, where n is the sample size. Under the assumption of constant population size over time, it follows that $\hat{\theta} = \tilde{\theta}$, which is equivalent to

$$S = \hat{\theta} \sum_{i=1}^{n-1} (1/i) . \quad (1)$$

If the observed number of variable sites S is significantly larger than $\hat{\theta}\Sigma_{i=1}^{n-1}(1/i)$, this indicates that an expansion may have affected the population (Tajima 1989). To test this, we computed the distributions of the number of variable sites by computer simulations, assuming constant population sizes over time and using the observed mean pairwise differences as estimates of θ. These distributions were then compared with observed numbers of variable sites.

In order to increase the sensitivity of the analyses, it was necessary to pool the 3 food-gathering and the 10 food-producing populations. The mean pairwise difference for the !Kung, Biaka, and Mbuti is 8.01, and 35 variable sites are observed among the 47 sequences. When equation (1) is used, with $\bar{d}$ as an estimate of θ, the expected number of variable sites is exactly 35. Hence, this test supports the view that the !Kung, Biaka, and Mbuti are more or less of constant size over the time covered by coalescence. For the food-producing populations, a mean pairwise difference of 7.9 and 118 variable sites among the 340 sequences are observed. However, 51 variable sites are expected. Since the observed number of 118 is outside the 95% confidence interval (i.e., 30–73), the notion that these groups have expanded their size is supported. Incidentally, if a random subsample of 47 sequences (the sample size of the !Kung, Biaka, and Mbuti) is taken from these groups, 59 variable sites are observed. This value is just larger than the critical value of 58 ($\alpha = .05$; one-sided test). Thus, even for a sample of the same size as the !Kung, Biaka, and Mbuti sample, the hypothesis of constant size for the other groups is rejected.

Most Recent Common Ancestors of the Populations

As shown above, the !Kung, Biaka, and Mbuti have been of constant size over a long time, whereas the other groups either are currently growing exponentially or have expanded in the past. For each of these three evolutionary scenarios it is possible to estimate the age of the most recent common ancestor of the sample. For the two first scenarios the method of Weiss and von Haeseler (1996) was applied, and for the sudden expansion model Rogers's (1995) formula for computing the time of an expansion was used. It can be seen in table 3 that, when the model with constant size is used, the coalescence times of the lineages in the !Kung, Biaka, and Mbuti are, on average, 45,000 years old. In contrast, when the models of exponential or sudden growth are used, the coalescences in the eastern African populations (Kikuyu, Somali, and Turkana) are slightly younger (30,000 years), whereas those in the western African populations are ~24,000 years old. The estimated times agree quite well with published ones (Sherry et al. 1994). One should be aware, however, that these estimates have a huge variance and that they are dependent on evolutionary rate estimates that are extremely tenuous.

Table 3

Estimates of Expected Times (Years) to the Most Recent Common Ancestor of the Population Samples, under Three Assumptions

Population	Constant Size[a]	Exponential Growth[b]	Sudden Expansion[c]
!Kung	40,000	34,000	4,000
Biaka	54,000	23,000	19,000
Mbuti	40,000	23,000	10,000
Kikuyu	116,000	29,000	35,000
Somali	103,000	27,000	34,000
Turkana	140,000	33,000	45,000
Fulbe	88,000	11,000	26,000
Hausa	81,000	21,000	NA
Songhai	91,000	47,000	36,000
Tuareg	94,000	25,000	26,000
Yoruba	101,000	21,000	32,000
Kanuri	94,000	33,000	29,000
Mandenka	79,000	9,000	21,000

NOTE.—The substitution rate used is .0025 per sequence and generation (Ward et al. 1991), and the generation time is 20 years. An underlined entry is a date that was calculated by using the model expected to be most appropriate for the population.

[a] The maximum-likelihood estimate of θ, given the number of observed variable sites, was computed. Based on this, the conditional probability distribution for the time to the most recent common ancestor was computed, and the expected time to coalescence for this distribution was estimated.

[b] Based on the current number of females (see table 1). The same computational procedure was employed as was used for "Constant Size."

[c] Estimate of expansion time if the model of "Pleistocene population explosion" (Rogers 1995) is adopted.

Discussion

Among the groups analyzed, the Mbuti, Biaka, and !Kung rely on hunting and food gathering for their subsistence, whereas the other groups practice agriculture and animal husbandry. The data suggest that the genetic diversity and demographic history of these two groups are dramatically different in at least four respects.

First, the 3 food-gathering groups have bumpy distri-

butions and a high proportion of identical lineages, whereas the 10 groups that rely on agriculture and/or cattle herding have bell-shaped distributions and few identical lineages (fig. 2). These patterns of distribution of pairwise sequence differences are reminiscent of the situation in other parts of the world. In India, the Mukri, a hunter-gatherer population, have a bumpy distribution, with 7.2% identical pairs, whereas the Havik, a food-producing population, display an essentially bell-shaped distribution, with 0.8% identical pairs (Mountain et al. 1995). Similarly, in Europe, three groups (British, Swiss, and Finns) have bell-shaped distributions, with <6% of pairs being identical (Pult et al. 1994), whereas the Saami, the only group in the region that until relatively recently was nonagriculturalist, have a rugged distribution, with 18% of the pairwise comparisons being identical (Sajantila et al. 1995). Thus, it seems to be a general pattern that food-producing populations have bell-shaped distributions of pairwise sequence differences, whereas food gatherers have rugged distributions and many identical sequences. An apparent exception to this are the Herero (Harpending et al. 1993), who were excluded from these analyses because of large amounts of undetermined nucleotide positions (Vigilant et al. 1991). A possible reason for this may be a recent population-size bottleneck that has affected the Herero (Harpending et al. 1993).

Second, the pattern of mitochondrial sequence variation in food-producing groups suggests that these populations have experienced an enlargement in size, whereas the food gatherers have been of constant population size for a long time. When this is taken into account, the time to coalescence is generally larger for the food-gathering groups than for the food-producing groups (table 3).

Third, the mitochondrial diversity in Africa indicates a substantial differentiation between food gatherers and food producers (table 2). This may not be altogether surprising. The transition from an economy based on hunting and gathering to one based on food production results in a drastic increase in population size (Ammerman and Cavalli-Sforza 1984), as well as in environmental changes. The latter are due both to the effects of food production per se—for example, clearing of forests and erosion by grazing—and to the environmental pressure caused by a larger number of individuals inhabiting an area. A group that has made this transition may therefore subsequently change its mode of food production but will be unlikely to return to an economy based on food gathering. Consequently, the groups that are hunter-gatherers today can be expected to have persisted as such for a long time. This is reflected in their extensive genetic differentiation from food-producing groups and from each other.

Fourth, the data indicate that the food-gathering groups have been isolated from each other over a long period, whereas food-producing groups differ much less from each other. This may be due to the common origin(s) of food-producing groups, which is associated with both the domestication of major crops (Ammerman and Cavalli-Sforza 1984) and a lower amount of genetic drift in the larger populations. It may also be due to more extensive female gene flow between food-producing groups, whose larger numbers and, in some cases, pastoralist lifestyle may facilitate contacts over large geographic distances. Further work in other African groups, as well as elsewhere, will show whether these differences between food-producing and food-gathering groups are typical of all or most such groups.

Acknowledgments

We thank the volunteers for participating in the study. Invaluable logistic support was provided by clinics and hospitals in Kenya (Turkana District Hospital, Lodwar; Maralal District Hospital; Isiolo District Hospital; Provincial General Hospital Nakuru; Sinna Medical Centre, Nairobi; and South Nyanza District Hospital, Kisumu) and Nigeria (Hopital National, Niamey; Akintunde Memorial Medical Centre, Saki; Baptist Medical Hospital, Saki; UDUTH, Sokoto; and UMTH, Maiduguri), as well as by the International Livestock Centre for Africa, in Nairobi. We thank the following persons for help and assistance: Dr. Abba Kyari Di, Prof. Idris Mohammed, Dr. Simon Obi, Prof. Chikwen, Dr. V. Brahasuriya, Dr. Jose P. Ambe, Dr. Simon Nwachuwa, Lawrence Ewesit, George Ochieng Otieno, Bashir Farah, Adirahman Abdullah Gacheru Muruthi, Obadish Nfunga, Michael Loter Kiyonga, Chaibou Tawaye, Dr. M. Akintunde, Dr. Amadou Sako, Dr. Hamma Soumana, and, especially, Usman Aliu Wakili the Seriki Fulani (Saki Idedapo Local Government Area), Sister Connie Hall (Mekaneyessus), Mohammed Hada, Mohammed Ahnou Abbou, Efe Efiawore, David Laur, Joseph Owino Jung'a, Jean-Romain da Silveira, Dr. E. Rege, Dr. Waudyallew Mulatu, Eshetu Zerihum, Dr. Jimmy Smith, Jonathon Ekuwuruke, Alhaji Nata'ala Sambo, Dr. Doug Little, Dr. Kwaku Agyemang, Dr. Mark Powell, Dr. Peter Lawrence, Dr. D. Penny, and F. Althaus. The Deutsche Forschungsgemeinschaft and the Deutsche Akademische Austauschdienst provided financial support.

References

Ammerman AJ, Cavalli-Sforza LL (1984) The Neolithic transition and the genetics of populations in Europe. Princeton University Press, Princeton, NJ

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290: 457–465

Batista O, Kolman CJ, Bermingham E (1995) Mitochondrial DNA diversity in the Kuna Amerinds of Panama. Hum Mol Genet 4:921–929

Bertranpetit J, Sala J, Calafell F, Underhill P, Moral P, Comas

D (1995) Human mitochondrial DNA variation and the origin of the Basques. Ann Hum Genet 59:63–81

Cabot EL (1993) ESEE: the eyeball sequence editor, version 3.0. University of British Columbia, Burnaby

Cann LS, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. Nature 325:31–36

DiRienzo A, Wilson AC (1991) Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc Natl Acad Sci USA 88:1597–1601

Felsenstein J (1990) PHYLIP, version 3.5p. University of Washington, Seattle

Graven L, Passarino G, Semino O, Boursot P, Santachiara-Benerecetti S, Langaney A, Excoffier L (1995) Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample. Mol Biol Evol 12:334–345

Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. Curr Anthropol 34:483–496

Hedges SB, Kumar S, Tamura K, Stoneking M (1992) Human origins and mitochondrial DNA sequences. Science 255: 737–739

Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N (1995) Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc Natl Acad Sci USA 92:532–536

Kolman CJ, Bermingham E, Cooke R, Ward RH, Arias TD, Guionneau-Sinclair F (1995) Reduced mtDNA diversity in the Ngöbé Amerinds of Panamá. Genetics 140:275–283

Maddison DR, Ruvolo M, Swofford DL (1992) Geographic origins of human mitochondrial DNA phylogenetic inference from control region sequences. Syst Biol 41:111–124

Marjoram P, Donnelly P (1994) Pairwise comparisons of mitochondrial DNA sequences in subdivided populations and implications for early human evolution. Genetics 136:673–683

Mountain JL, Hebert JM, Bhattacharyya S, Underhill PA, Ottolenghi C, Gadgil M, Cavalli-Sforza LL (1995) Demographic history of India and mtDNA-sequence diversity. Am J Hum Genet 56:979–992

Murdock GP (1959) Africa: its peoples and their culture and history. McGraw Hill, New York

Nee S, Holmes EC, Rambaut A, Harvey PH (1995) Inferring population history from molecular phylogenies. Philos Trans R Soc Lond [Biol] 349:25–31

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Penny D, Steel M, Waddell PJ, Hendy MD (1995) Improved analyses of human mtDNA sequences support a recent African origin of *Homo sapiens*. Mol Biol Evol 12:863–882

Piercy R, Sullivan KM, Benson N, Gill P (1993) The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. Int J Legal Med 106:85–90

Pult I, Sajantila A, Simanainen J, Georgiev O, Schaffner W, Pääbo S (1994) Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. Biol Chem Hoppe-Seyler 375:837–840

Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. Evolution 49:608–615

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise differences. Mol Biol Evol 9:552–569

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sajantila A, Lahermo P, Anttinen T, Lukka M, Sistonen P, Savontaus M-L, Aula P, et al (1995) Genes and languages in Europe: an analysis of mitochondrial lineages. Genome Res 5:42–52

Santos M, Ward RH, Barrantes R (1994) mtDNA variation in the Chibcha Amerindian Huetar from Costa Rica. Hum Biol 66:963–977

Sherry ST, Rogers AR, Harpending H, Soodyall H, Jenkins T, Stoneking M (1994) Mismatch distributions of mtDNA reveal recent human population expansions. Hum Biol 66: 761–776

Shields GF, Schmiechen AM, Frazier BL, Redd A, Voevoda MI, Reed JK, Ward RH (1993) MtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. Am J Hum Genet 53:549–562

Slatkin M, Hudson RR (1991) Pairwise comparisons in stable and exponentially growing populations. Genetics 129:555–562

Tajima F (1989) The effect of change in population size on DNA polymorphism. Genetics 123:597–601

Templeton AR (1992) Human origins and mitochondrial DNA sequences. Science 255:737

——— (1993) The "Eve" hypothesis: a genetic critique and reanalysis. Am Anthropol 95:51–72

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, et al (1993*a*) Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53:563–590

Torroni A, Sukernik RI, Schurr TG, Starikovskaya YB, Cabell MF, Crawford MH, Comuzzie AG, et al (1993*b*) mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. Am J Hum Genet 53:591–608

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. Science 253:1503–1507

Ward RH, Frazier BL, Dew K, Pääbo S (1991) A single North-American tribal group contains extensive mitochondrial diversity. Proc Natl Acad Sci USA 88:8720–8724

Ward RH, Redd A, Valencia D, Frazier B, Pääbo S (1993) Genetic and linguistic differentiation in the Americas. Proc Natl Acad Sci USA 90:10663–10667

Weiss G, von Haeseler A (1996) How many Y's? Science 272: 1359–1360

Zischler H, Geisert H, von Haeseler A, Pääbo S (1995) A nuclear "fossil" of the mitochondrial D-loop and the origin of modern humans. Nature 378:489–492